A TEST OF THE ATTAINMENT OF FIRST-YEAR HIGH-SCHOOL STUDENTS IN ALGEBRA[1]

WALTER S. MONROE
Kansas State Normal School, Emporia, Kansas

The algebra taught in our secondary schools includes two distinct types of subject-matter. One type is made up of the formal processes which consist of manipulations of symbols according to certain rules that have been agreed upon by those who developed the subject. For example, when one solves the equation $6x-x^2+4=2(x-3)(x+2)$, he manipulates the symbols of that equation according to a plan of procedure which has been found satisfactory. First the equation is put in the form, $6x-x^2+4=2x^2-2x-12$. Then it is changed to the form, $3x^2-8x-16=0$. This equation, which is recognized as a quadratic, may now be solved by manipulating the symbols according to one of several plans of procedure. For example, if one forms the product of 4 times 3 times 16 which is 192, adds this to the product of 8 times 8, takes the square root of the sum, 256, which is 16, and finally adds 16 to 8, and subtracts 16 from 8, this sum and difference when divided by 6 are the roots of the above equation.

It is with this type of algebraic subject-matter that we are concerned in this investigation. The formal processes are necessary prerequisites for the solution of practical problems. After the conditions of a problem are expressed in the form of an equation, it is necessary that the symbols of that equation be manipulated so that the values desired can be obtained.

Proficiency in algebra demands that one's knowledge of the formal processes be reduced to the plane of habit and be capable of automatic application. This is just what is demanded in arithmetic. For example, knowledge of the fundamental combinations of the multiplication table must become a matter of habit and the

[1] The writer is indebted to Ernest E. Wellemeyer, Oklahoma City High School, for scoring a large number of the papers and for assistance in giving the tests.

correct product be produced without the higher thought processes being involved. Speed and accuracy, not ability to think or reason out the products, are demanded after the very earliest stages of the child's education. So it is in algebra. The demand so far as concerns the formal side of algebra is for speed and accuracy in performing additions, subtractions, multiplications, and divisions, in removing parentheses, in clearing equations of fractions, in completing squares, etc. It is true that many teachers believe that, when a student is learning the rules for manipulating symbols, the best results are secured when he develops the rules for himself, i.e., when the rules are for him the product of his own thinking. But even those who make this contention agree that a little later the knowledge should be capable of automatic application, at least for those processes which possess fundamental importance. The advantage of reducing the formal processes to the plan of habit is clearly evident when it is realized that the symbols of a given type of example are always to be manipulated in the same way.

The speed and accuracy with which a student performs such formal processes as the multiplication of one factor by another, or transposing terms, or clearing an equation of fractions can easily be measured. And, since these two qualities define the ability to perform those processes which have been reduced to the plane of habit, we may in this way obtain a measure of these algebraic abilities of a student.

It has been the purpose of the writer to devise a set of tests which could be used to measure algebraic abilities and to determine the standards which should be attained in these abilities. In order to select types of examples which are fundamentally important, simple equations with numerical coefficients and numerical denominators were taken, and the tests limited to the distinct types of examples involved in solving such equations. The several types of examples were combined in test F which is reproduced here. The other tests were as follows:

Test A, to multiply $\pm a(\pm bx \pm c)$ where a, b, and c in no case were greater than 9, and in no case were all positive.

Test B, to reduce fractions to a common denominator.

Test C, to solve equations of the type $\pm ax = \pm b$.

ALGEBRA TEST F. SIMPLE EQUATIONS

Name________________________________ Date________________________

Find the value of x in the following equations. Solve the examples on another sheet but be sure to hand in all your work. The solution of an equation should appear in the following form:

$$\frac{-6x-1}{8}-(-4x-8)=\frac{5(3x+4)}{6}$$

$$-18x+3+96x+192=60x+80$$

$$-18x+96x-60x=80-3-192$$

$$18x=-115$$

$$x=-\frac{115}{18}$$

You may leave the value of x in fractional form as above. Do *not* check your work. Work the examples in order as numbered. Do as many as you can in the time allowed, but remember that accuracy counts as well as speed.

1. $-\frac{3x-2}{4}=\frac{x+2}{6}$
2. $-\frac{-5+4x}{3}+2=-4(3x+2)$
3. $3(1-7x)-7-7(-4x+5)=-4(9-4x)$
4. $1-\frac{x+5}{6}=\frac{-4x+9}{8}$
5. $\frac{1-x}{4}-\frac{-3x-4}{8}=\frac{6x+7}{6}$
6. $\frac{4x+5}{15}-\frac{2x+3}{21}=0$
7. $-\frac{3(2x-5)}{5}+\frac{-2(3x+4)}{7}=0$
8. $2-\frac{-5(3+2x)}{6}=5x-4$
9. $7x-5\frac{6x-11}{3}=12$
10. $5+\frac{3x+1}{8}-\frac{-1+2x}{12}=0$
11. $\frac{-4x-3}{6}=8-\frac{4(x-3)}{9}$
12. $4x-9\frac{-4x+7}{8}=3+5\frac{3x-2}{14}$
13. $\frac{2x+3}{10}-\frac{1+2x}{12}+\frac{-4x-5}{15}=1$

Test D, to transpose terms in equations such as $4x-6+5=7x-4-2$.

Test E, to collect terms in expressions such as $-5x+6x-11x+8-3-9$.

In constructing tests B, C, D, and E the examples in each test were made similar to the corresponding step in the solution of the corresponding example in test F. For example, the solution of the first example is as follows:

$$-\frac{3x-2}{4}=\frac{x+2}{6}$$
$$-18x+12=4x+8$$
$$-18x-4x=-12+8$$
$$-22x=-4$$
$$x=4/22$$

The first examples in the other four tests are:

Test B, reduce to a common denominator $-\frac{7x-2}{6}+\frac{x+1}{8}$

Test C, solve $-15x=-4$, answer to be left in the form $x=4/15$.

Test D, transpose the terms containing x to the left-hand side of the equality sign and those not containing x to the right-hand side in $-4x+5=3x-5$.

Test E, collect the terms in $-7x-3x-6+4$.

A mimeographed copy of each test was given to each student. The tests were given by the regular teachers, who were requested to follow printed directions, to twelve classes of first-year students in two city high schools in March, 1914. In both schools, the students who were not doing satisfactory work at the middle of the year were placed in a beginning class. These students were not tested. Owing to a misunderstanding, test B was not given to four classes. The time allotment was as follows:

Test A, 2 minutes	Test D, 2 minutes
Test B, 3 "	Test E, 3 "
Test C, 1 "	Test F, 12 "

The first five were given on one day, and test F on the following day.

In scoring the tests, an example was not counted as attempted unless it had been completed. Thus if a pupil had nearly com-

pleted the tenth example in test F, he was counted as having attempted only 9. However, the method used in computing the averages makes an allowance for this apparent unfairness. After the papers of a class had been scored, they were distributed, first according to the number of examples attempted, and second according to the number of examples right. A statement of the number of pupils who attempted 5 examples, 7 examples, etc., is called the *distribution* of the class according to the number of examples attempted. The total distributions by tests are given in Table I. From these distributions the averages and the average deviations of the group from this average have been computed by the "short" method.

A study of Table I shows that some students worked very rapidly and a few worked very slowly. One student performed 34 multiplications in three minutes while in the same time another pupil performed only 3 multiplications. These are the extreme cases, but the remainder of the group are quite widely distributed. The average number attempted is 17.3, which is almost exactly half of the maximum number. There is a large number of students who worked more rapidly than this average and also many who worked less rapidly. In the other tests the number of examples solved is less and hence the range of distribution is less, but the same general type prevails. The distribution of students according to the number of examples right, with the exception of test A, shows a relatively large number of students who failed to get a single example right.

Similar distributions of students have been shown by other tests, and it is not surprising that a considerable range of variation should exist when specific algebraic abilities are measured. However, the magnitude of the range suggests that those students who work slowly have not been given just the type of training which they need. Certainly the student who can perform only three multiplications, or even ten multiplications, in two minutes needs a type of training essentially different from that demanded by a student who can perform twenty multiplications in the same time. To subject them to the same type of training is probably to give neither the training which he needs.

In Table II the class averages and average deviations are given. One notes immediately a marked difference in averages of the

TABLE I

TOTAL DISTRIBUTIONS

No. Examples	Attempts						Right					
	Test A	B	C	D	E	F	A	B	C	D	E	F
34	1	....	....	....	....	....	....	....	....	....	....	....
33	1	....	....	....	....	....	....	....	....	....	....	....
32	....	....	....	....	....	....	1	....	....	....	....	....
31	....	....	....	....	....	....	....	....	....	....	....	....
30	1	....	....	....	....	....	1	....	....	....	....	....
29	1	....	....	....	....	....	....	....	....	....	....	....
28	1	....	....	....	....	....	....	....	....	....	....	....
27	5	....	....	....	....	....	2	....	....	....	....	....
26	4	....	....	....	....	....	3	....	....	....	....	....
25	4	....	....	....	....	....	3	....	....	....	....	....
24	7	....	....	....	....	....	4	....	....	....	....	....
23	13	....	....	....	....	....	10	....	....	....	....	....
22	15	....	....	....	....	....	8	....	....	....	....	....
21	12	....	....	....	....	....	10	....	....	....	....	....
20	8	....	....	....	....	....	13	....	....	....	....	....
19	15	....	....	....	....	....	9	....	....	....	....	....
18	26	....	5	....	....	....	17	....	1	....	....	....
17	27	....	2	....	....	....	16	....	1	....	....	....
16	29	....	3	....	....	....	24	....	3	....	....	....
15	33	....	7	....	....	....	26	....	5	....	....	....
14	14	....	12	....	3	....	28	....	8	....	....	....
13	23	....	21	5	14	1	21	....	13	3	2	....
12	10	....	20	26	55	1	12	....	11	5	7	....
11	9	....	21	37	58	3	8	....	17	20	12	....
10	7	....	21	45	40	8	13	....	16	32	32	....
9	2	2	35	46	39	19	7	....	23	35	32	....
8	3	2	25	37	25	37	12	...	17	39	31	3
7	1	13	24	26	16	53	5	....	15	25	33	4
6	2	40	32	21	12	61	6	1	32	24	31	7
5	....	57	24	19	9	39	4	5	14	26	23	17
4	....	60	14	5	3	20	2	10	14	12	22	22
3	1	14	5	3	1	5	4	34	13	9	17	26
2	....	8	1	3	....	2	2	36	9	2	8	41
1	....	1	3	1	....	....	2	44	19	11	6	55
0	....	....	....	....	....	....	2	67	44	31	19	74
Total	275	197	275	274	275	249	275	197	275	274	275	249
Av.	17.3	5.4	9.5	9.3	10.4	7.1	15.3	2.0	6.2	7.1	7.0	2.4
A.D.	3.5	1.0	2.9	1.8	1.7	1.0	4.3	1.3	3.9	2.8	2.6	1.6

several classes. In test A, class 8 has an average of 21.5 examples attempted while class 5 has an average of 12.2. Since these classes are from the same school, the difference in ability is due to the

difference in the instruction or in the native ability of the students. Classes 9, 10, 11, and 12 were taught by the same teacher. They show rather marked differences in both averages and deviations. Class 12 has a high average and one of the smallest deviations.

TABLE II

CLASS AVERAGES AND AVERAGE DEVIATIONS

Class	Examples Attempted											
	Test A		B		C		D		E		F	
	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.
1.....	17.7	3.0	5.2	1.0	11.5	2.6	10.5	1.3	10.3	1.5	7.4	1.3
2.....	18.8	2.7	5.6	0.9	10.0	2.1	10.2	1.4	10.8	1.9	7.7	1.2
3.....	17.5	2.8	5.5	1.1	7.5	2.1	10.4	1.5	11.3	1.3	7.6	1.4
4.....	16.4	3.5	5.0	1.0	8.9	2.3	9.8	1.3	10.9	1.1	7.0	1.0
5.....	12.2	3.4	5.2	0.8	9.9	4.0	8.8	1.7	11.0	1.9	7.3	1.3
6.....	16.7	2.4	5.3	0.9	8.5	2.3	9.6	1.1	10.6	1.7	6.2	1.1
7.....	18.4	3.4	5.2	0.9	8.0	2.0	9.8	1.0	10.2	1.9	6.7	1.0
8.....	21.5	4.4	5.7	1.5	11.0	2.7	10.5	2.0	10.6	2.0	6.1	2.4
9.....	17.4	2.7	.....	.....	9.0	3.4	6.5	1.7	9.4	1.6	6.2	1.7
10....	17.8	4.0	.....	.....	10.1	2.4	7.0	2.3	9.4	2.6	6.8	1.4
11....	19.7	4.6	.....	.....	8.2	2.7	7.7	1.5	9.5	1.5	7.1	1.4
12....	20.6	2.6	.....	.....	10.0	4.0	8.7	1.3	9.8	2.0	6.2	0.8

Class	Examples Right											
	Test A		B		C		D		E		F	
	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.	Av.	A.D.
1.....	15.9	3.8	2.6	1.1	9.2	3.9	9.3	1.8	7.9	2.1	3.3	1.6
2.....	15.5	4.5	2.0	2.0	9.2	3.0	8.1	1.6	6.0	3.1	2.4	1.3
3.....	16.0	3.7	1.2	1.1	5.8	2.9	8.2	2.3	7.5	2.6	2.4	1.7
4.....	14.4	3.8	1.4	1.0	6.6	3.3	7.8	2.1	7.5	2.5	2.1	1.3
5.....	9.2	5.2	1.8	1.1	6.3	4.7	6.2	2.5	6.9	1.1	2.3	1.3
6.....	13.5	3.3	2.2	1.0	6.7	3.1	8.0	2.1	7.6	2.6	2.7	1.5
7.....	15.6	3.8	2.2	1.3	6.6	2.5	8.5	1.3	7.9	2.6	3.6	1.9
8.....	19.3	4.5	1.3	1.0	7.0	4.0	9.8	2.3	6.9	3.3	1.9	1.6
9.....	14.3	3.0	.....	.....	4.2	4.2	2.2	2.5	5.9	1.6	1.9	1.1
10....	14.8	2.8	.....	.....	8.3	2.0	3.9	2.6	6.0	3.0	1.9	1.4
11....	16.8	4.4	.....	.....	4.0	3.6	4.8	2.8	6.3	2.1	2.0	1.4
12....	18.8	4.8	.....	.....	5.8	4.5	7.3	2.2	7.2	1.9	1.9	0.9

The average of class 10 is nearly three examples less than and the average deviation is nearly double that of class 12. This condition indicates the relative efficiency of the teaching in the two classes. Approximately the same method was followed in

both, but evidently it was much more successful in class 12 than in class 10. The members of class 12 exhibit a high degree of ability to multiply and in this there is a high degree of uniformity. In this case the instruction has been well suited to the various members of the class. In the case of class 10, it is evident that the instruction was not so well suited to all members of the class. It was a member of class 10 who made the highest score of examples attempted in the entire group, but 60 per cent of the class made scores less than 17. The method of instruction developed a relatively low degree of ability in these students. It is probable that a different plan of procedure would have given more desirable results.

In tables not reproduced here, it is shown that there is no evidence in support of the frequently expressed belief that girls excel boys in facility in algebra, for in three of the tests the average number attempted by the boys is higher than the corresponding average for the girls, and in the other three tests the conditions are reversed. In every test the boys showed superior accuracy.

The class averages represent the measure of the average ability of the class. For example, the average ability of class 2 in test F is represented by the class average of 7.7 examples attempted. As a tentative standard of this ability we may take the average of the total group, which is 7.1. Comparing the average of class 2 we see that it is 0.6 of an example above the standard; that is, the average of the class in this particular ability is superior to the standard. On the other hand, comparing the average of class 8 with the standard, we find that it is inferior.

The averages and the average derivations from these tests may be used as standards for the purpose of comparing other classes with these, but the writer thinks that they are not satisfactory as permanent standards. This is particularly true when the accuracy of the work is considered. For example, in test F only approximately one example out of three was correct. The percentage of examples right was larger in the other tests, but it must be remembered that multiplication, transposition, etc., are useful only when they are combined as in test F. Hence, we may interpret the results of this test as follows. After six months of instruction the

average first-year high-school student can solve correctly only one out of three simple equations which he attempts when working under the conditions of this test and when the equations possess the difficulty of those in test F.

The correlations of the relative rank of the several classes on the basis of the number of examples attempted except in test B are as shown in Table III.

TABLE III

	Test A	C	D	E	F
Coefficient of correlation with rank in test A	....	.29	−.21	−.38	−.12
Coefficient of correlation with rank in test C	.29	....	.18	−.21	−.11
Coefficient of correlation with rank in test D	.21	.18	....	.64	.30
Coefficient of correlation with rank in test E	−.38	−.21	.64	....	.61
Coefficient of correlation with rank in test F	−.12	−.11	.30	.61	....

Perfect correlation would be expressed by a coefficient of 1.00. With two exceptions the coefficients are relatively small. The meaning of this is that each of the tests calls for a different ability.

The number of students who failed to get a single example right is worthy of comment. In test B, 67 students, or 34 per cent, failed to get a single example right. A study of this group shows that the average number of examples attempted was 5.1 and that the average deviation was 1.12. These are almost identical with the average and average deviation for the entire group. This suggests that accuracy is not dependent upon the speed with which a student works. This question was studied in more detail for the entire group.

Table IV shows the distribution of the 250 students who took test F according to the number of examples attempted and the number of examples right. This table is read as follows: Of the eight students attempting 10 examples, one had 8 right, one 7 right, one 6 right, two 5 right, two 1 right, and one none right. From such a distribution it is easy to calculate the total number of examples attempted by those attempting 2 problems, by those attempting

3 problems, etc., and also the total number right by each group. From these two totals the percentage of examples right can be calculated for each group. This percentage is given in the last column of Table IV and expresses the accuracy with which each group worked in test F. For example, those students who attempted 10 examples worked with an average accuracy of 41 per cent, those attempting 9 examples worked with an average accuracy of 34 per cent, etc. A comparison of the percentage of examples right with the number attempted by the group immediately suggests a very close correlation between the two.

TABLE IV

TEST F

Rights / Attempts	8	7	6	5	4	3	2	1	0	Total Examples Attempted	Total Right	Percentage of Examples Right
13......	1	.....	.....	.....	.....	.....	.....	.....	.....	13	8	61
12......	.....	.....	.....	.....	.....	.....	.....	.....	1	12	0	00
11......	.....	1	.....	1	.....	1	.....	.....	.....	33	15	45
10......	1	1	1	2	.....	.....	.....	2	1	80	33	41
9.......	.....	2	1	3	2	2	2	6	1	171	59	34
8.......	1	.....	1	3	6	5	10	5	7	304	85	28
7.......	.....	.....	3	5	7	7	4	14	13	371	114	31
6.......	.....	.....	1	3	6	7	14	12	18	366	96	26
5.......	.....	.....	.....	.....	1	5	7	12	14	195	45	23
4.......	.....	.....	.....	.....	.....	.....	3	3	12	72	9	13
3.......	.....	.....	.....	.....	.....	.....	.....	2	5	21	2	10
2.......	.....	.....	.....	.....	.....	.....	.....	1	1	4	2	50
Total..	3	4	7	17	22	27	41	56	73	....	.....	.....
Average	10.8	10.3	8.2	8.45	7.6	7.35	6.8	7.0	6.3	....	.....	.....

In order to be certain that this relation was not due to the accidental combination of several classes, a similar distribution was made for each class and the percentage right was calculated for each group. The result arrived at by this procedure shows that, although the degree of correlation is not always high, the general tendency is for a relatively high degree of accuracy to accompany the more rapid work.

Table V gives the relation between the number of examples attempted and the percentage right for each test except test A. A

study of Table V shows that there is a very high correlation between the number of examples attempted and the accuracy with which these examples were worked, except in test B and test F. The low coefficient of correlation in test F is due almost entirely to two groups of students, those attempting 12 examples and those attempting 2 examples. These two groups involve a total of 3 students out of 250. For this reason the correlation is much closer than is indicated by the computed coefficient. The same conditions exist with respect to test B.

TABLE V

SHOWING RELATION BETWEEN THE NUMBER OF EXAMPLES ATTEMPTED AND THE PERCENTAGE RIGHT

No. Examples Attempted	Total for All Classes				
	Test B	C	D	E	F
18	....	69	....	....	....
17	....	85	....	....	....
16	....	85	....	....	....
15	....	93	....	....	....
14	....	71	....	88	....
13	....	75	89	71	61
12	....	83	85	71	00
11	....	71	83	70	45
10	....	80	84	71	41
9	17	72	79	61	34
8	44	57	71	56	28
7	33	73	73	48	31
6	29	63	47	49	26
5	34	55	31	40	23
4	30	57	32	17	13
3	29	11	00	....	10
2	19	00	17	....	50
1	00	00	00	....	....
Coefficient of correlation	.42	.80	.98	.98	.29

This high degree of correlation means that of these high-school students, the ones who work rapidly work with a relatively higher degree of accuracy than those who work more slowly. This is contrary to the popular belief that those who work slowly work with greater care and accuracy. Evidently students who work rapidly may be just as accurate as those who work slowly. In fact, the data here show that they are superior in accuracy. Thus accuracy must depend upon some factor other than the speed with

which the student works. This being true, it seems reasonable to expect that accuracy can be developed independently of speed, and that its development is not to be secured just by cautioning the student to work more slowly.

Table VI shows a classification of the errors which were made in the several tests. Using the wrong sign is a prolific source of inaccuracy. Even in such simple examples as in test A, the number of errors is quite large. The errors under the head of arithmetic in test C are of the type of writing $x=5/8$ instead of $x=8/5$, as the value of x in $5x=8$. Errors of the last four types involve primarily attention and not knowledge. For example, in test D, when a term is copied $7x$ instead of $5x$, the student has failed to give the requisite degree of attention to the work.

TABLE VI

CLASSIFICATION OF THE ERRORS

	Test A	B	C	D	E	F
Mistakes in sign	436	263	184	295	387	739
Mistakes in the common denominator or in its use	......	167	......	......	......	371
Mistakes in arithmetic	143	249	498	......	596	391
Mistakes in copying	......	......	......	382	63	82
One term of binomial not multiplied	29	......	......	......	......	16
A term neglected	......	103	......	302	86	26
x omitted	117	19	......	53	80	16
Incomplete as $-x=5$	......	......	38	......	......	26

These facts indicate that the errors made in these tests were due to inattention on the part of the students or to the failure to cause the students to reduce these manipulations of symbols to the plane of habit. Doubtless every student possessed sufficient knowledge, but it failed to function because he was not sufficiently attentive or the knowledge was not capable of automatic application.

If our analysis is correct, improvements may be made in two ways. First, the student is attentive to what appeals to him as worth while, either because of some present interest or because of the realization of future usefulness. Thus the teacher who succeeds in causing his pupils to take an interest in solving equations

will at the same time secure increased accuracy because of the increased attention. This interest will be more permanent if it is based upon the appreciation of the equation as a practical tool.

The other source of improvement is to make automatic each type of manipulation which is required. This requires systematic practice upon each detail. For example, in test C there are four subtypes, $ax=b$, $-ax=b$, $ax=-b$, $-ax=-b$. It is necessary to drill systematically upon each of these subtypes. The drill not only should be systematic but in addition it should be intelligently applied. When the students are drilled upon a type of example, both the individual student and the teacher should know in just what subtypes he is making mistakes. In cases where the student does not soon remedy the defect thus revealed, a detailed diagnosis should be made by the teacher and appropriate remedies prescribed.

This work was undertaken as an attempt to devise a means for measuring a few of the important algebraic abilities and to determine the standard of these abilities for first-year high-school students. A valuable by-product has been the analysis of existing conditions. The averages which are given should be considered only as tentative standards. With this in mind, they will be useful for purposes of comparison.